



1

SEQUENCE LISTING

RECEIVED

JUL 20 2001

TECH CENTER 1600/2900

<10> BECKWITH, JONATHAN
ASLUND, FREDRIK
BESSETTE, PAUL H.
GEORGIU, GEORGE
RITZ, DANIEL
LIM, JACKIE EUN-AH

<120> COMPOSITIONS AND METHODS FOR PRODUCTION OF DISULFIDE
BOND CONTAINING PROTEINS IN HOST CELLS

<130> HMV-052.01

<140> 09/679,705
<141> 2000-10-05

<150> 60/157,770
<151> 1999-10-05

<150> 60/163,939
<151> 1999-11-08

<150> 60/166,044
<151> 1999-11-17

<160> 24

<170> PatentIn Ver. 2.1

<210> 1
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Illustrative
motif

<220>
<221> MOD_RES
<222> (2)..(3)
<223> Any amino acid

<400> 1
Cys Xaa Xaa Cys
1

<210> 2
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
variant motif

<400> 2
Cys Gly Pro Cys
1

<210> 3
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
variant motif

<400> 3
Cys Gly Ser Cys
1

<210> 4
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
variant motif

<400> 4
Cys Pro Tyr Cys
1

<210> 5
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
variant motif

<400> 5
Cys Pro His Cys
1

<210> 6
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
variant motif

<400> 6
Cys Gly His Cys
1

<210> 7
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
variant motif

<400> 7
Cys Gly Pro Ala
1

<210> 8
<211> 48
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)..(48)

<400> 8
tgg agc gtc ttc ttc ttc tac ccg gct gac ttt act ttc gta tgc ccg 48
Trp Ser Val Phe Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys Pro
1 5 10 15

<210> 9
<211> 16
<212> PRT
<213> Escherichia coli

<400> 9
Trp Ser Val Phe Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys Pro
1 5 10 15

<210> 10
<211> 51
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)..(51)

<400> 10
tgg agc gtc ttc ttc ttc ttc tac ccg gct gac ttt act ttc gta tgc 48
Trp Ser Val Phe Phe Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
1 5 10 15

ccg
Pro

51

<210> 11
<211> 17
<212> PRT
<213> Escherichia coli

<400> 11
Trp Ser Val Phe Phe Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
1 5 10 15

Pro

<210> 12
<211> 32
<212> PRT
<213> Escherichia coli

<400> 12
Arg Trp Ser Val Phe Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
1 5 10 15

Pro Thr Glu Leu Gly Asp Val Ala Asp His Tyr Glu Glu Leu Gln Lys
20 25 30

<210> 13
<211> 32
<212> PRT
<213> Salmonella typhi

<400> 13
Arg Trp Ser Val Phe Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
1 5 10 15

Pro Thr Glu Leu Gly Asp Val Ala Asp His Tyr Glu Glu Leu Gln Lys
20 25 30

<210> 14
<211> 32
<212> PRT
<213> Pseudomonas putida

<400> 14
Lys Trp Ser Val Val Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
1 5 10 15

Pro Thr Glu Leu Gly Asp Leu Ala Asp Asn Tyr Ala Glu Phe Gln Lys
20 25 30

<210> 15
<211> 32

<212> PRT

<213> *Staphylococcus mutans*

<400> 15

Lys Trp Ala Val Phe Cys Phe Tyr Pro Ala Asp Phe Ser Phe Val Cys
1 5 10 15

Pro Thr Glu Leu Gly Asp Leu Gln Glu Gln Tyr Ala Thr Leu Gln Ser
20 25 30

<210> 16

<211> 32

<212> PRT

<213> *Bacillus subtilis*

<400> 16

Gln Trp Ser Val Phe Cys Phe Tyr Pro Ala Asp Phe Ser Phe Val Cys
1 5 10 15

Pro Thr Glu Leu Glu Asp Leu Gln Glu Gln Tyr Ala Ala Leu Lys Glu
20 25 30

<210> 17

<211> 32

<212> PRT

<213> *Staphylococcus aureus*

<400> 17

Ser Trp Ser Val Val Cys Phe Tyr Pro Ala Asp Phe Ser Phe Val Cys
1 5 10 15

Pro Thr Glu Leu Glu Asp Leu Gln Asn Gln Tyr Glu Glu Leu Gln Lys
20 25 30

<210> 18

<211> 32

<212> PRT

<213> *Treponema pallidum*

<400> 18

Ser Trp Ala Val Phe Met Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
1 5 10 15

Pro Thr Glu Leu Ala Asp Leu Ala Arg Val Tyr Pro Ser Phe Val Glu
20 25 30

<210> 19

<211> 32

<212> PRT

<213> *Aquifex aeolicus*

<400> 19

Lys Trp Val Ile Leu Phe Phe Tyr Pro Ala Asp Tyr Thr Phe Val Cys
1 5 10 15

<400> 21																
aagggtagtt	cagattacac	ggtcacctgg	aaagggggcc	attttacttt	ttatgcgctg											60
gcggtgcaaa	gttcacaaag	ttgtcttacg	aaggttgtaa	ggtaaaactt	atcgatttga											120
taatggaaac	gcattaccgg	aatcggcaaa	aattggttac	cttacatctc	atcgaaaaca											180
cggaggaagt	atag	atg	tcc	ttg	att	aac	acc	aaa	att	aaa	cct	ttt	aaa			230
		Met	Ser	Leu	Ile	Asn	Thr	Lys	Ile	Lys	Pro	Phe	Lys			
		1				5					10					
aac	cag	gca	ttc	aaa	aac	ggc	gaa	ttc	atc	gaa	atc	acc	gaa	aaa	gat	278
Asn	Gln	Ala	Phe	Lys	Asn	Gly	Glu	Phe	Ile	Glu	Ile	Thr	Glu	Lys	Asp	
		15					20					25				
acc	gaa	ggc	cgc	tggt	agc	gtc	ttc	ttc	ttc	tac	ccg	gct	gac	ttt	act	326
Thr	Glu	Gly	Arg	Trp	Ser	Val	Phe	Phe	Phe	Tyr	Pro	Ala	Asp	Phe	Thr	
	30					35					40					
ttc	gta	tgc	ccg	acc	gaa	ctg	ggc	gac	gtt	gct	gac	cac	tac	gaa	gaa	374
Phe	Val	Cys	Pro	Thr	Glu	Leu	Gly	Asp	Val	Ala	Asp	His	Tyr	Glu	Glu	
	45				50				55					60		
ctg	cag	aaa	ctg	ggc	gta	gac	gta	tac	gca	gta	tct	acc	gat	act	cac	422
Leu	Gln	Lys	Leu	Gly	Val	Asp	Val	Tyr	Ala	Val	Ser	Thr	Asp	Thr	His	
				65					70					75		
ttc	acc	cac	aaa	gca	tggt	cac	agc	agc	tct	gaa	acc	atc	gct	aaa	atc	470
Phe	Thr	His	Lys	Ala	Trp	His	Ser	Ser	Ser	Glu	Thr	Ile	Ala	Lys	Ile	
			80					85					90			

aaa tat gcg atg atc ggc gac ccg act ggc gcc ctg acc cgt aac ttc	518
Lys Tyr Ala Met Ile Gly Asp Pro Thr Gly Ala Leu Thr Arg Asn Phe	
95 100 105	
gac aac atg cgt gaa gat gaa ggt ctg gct gac cgt gcg acc ttc gtt	566
Asp Asn Met Arg Glu Asp Glu Gly Leu Ala Asp Arg Ala Thr Phe Val	
110 115 120	
gtt gac ccg cag ggt atc atc cag gca atc gaa gtt acc gct gaa ggc	614
Val Asp Pro Gln Gly Ile Ile Gln Ala Ile Glu Val Thr Ala Glu Gly	
125 130 135 140	
att ggc cgt gac gcg tct gac ctg ctg cgt aaa atc aaa gca gca cag	662
Ile Gly Arg Asp Ala Ser Asp Leu Leu Arg Lys Ile Lys Ala Ala Gln	
145 150 155	
tac gta gct tct cac cca ggt gaa gtt tgc ccg gct aaa tgg aaa gaa	710
Tyr Val Ala Ser His Pro Gly Glu Val Cys Pro Ala Lys Trp Lys Glu	
160 165 170	
ggt gaa gca act ctg gct ccg tct ctg gac ctg gtt ggt aaa atc taa	758
Gly Glu Ala Thr Leu Ala Pro Ser Leu Asp Leu Val Gly Lys Ile	
175 180 185	
atttccttag tcttttcacgc atagcggcgt tgcgtcgccc gctcaccggt tcaacttactt	818
gtgtaagctc ccggggattc acagctagcg ccttgctctg acgcgaaata cttcggaat	878
tcacctaatt cttcggtgc tgcggcgcat tttcttcccc gcaccatgat gcaagctgca	938
tccaggtagc cgcagaggcc gcttgcatga tgatgtttaa gagcccagga gataaacatg	998
ctcgacacaa atatgaaaac tcaactcaag gcttaccttg agaaattgac caagcctggt	1058
gagttaattg ccacgctgga tgacagcgct aaatcggcag aaatcaagga actggtggct	1118
gaaatcgcag aactgtcaga caaagtcacc tttaaagaag ataacagctt gccggtgcgt	1178
aagccgtctt tctgatcac caaccaggt tccaaccagg ggccacgttt tgcaggctct	1238
ccgctgggcc acgagttcac ctgcgtggtg ctggcggtgc tgtggaccgg tggatcatccg	1298
tcgaaagaag cgcagtctct gctggagcag attcgccata ttgacggtga ttttgaattc	1358
gaaacctatt actcgtcttc ttgccacaac tgcccggacg tgggtgcaggc gctgaacctg	1418
atgagcgtac tgaaccgcg catcaagcac actgcaattg acggcggcac cttccagaac	1478
gaaat	1483

<210> 22

<211> 187

<212> PRT

<213> Escherichia coli

<400> 22

```

Met Ser Leu Ile Asn Thr Lys Ile Lys Pro Phe Lys Asn Gln Ala Phe
 1           5           10           15

Lys Asn Gly Glu Phe Ile Glu Ile Thr Glu Lys Asp Thr Glu Gly Arg
          20           25           30

Trp Ser Val Phe Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys Pro
          35           40           45

Thr Glu Leu Gly Asp Val Ala Asp His Tyr Glu Glu Leu Gln Lys Leu
 50           55           60

Gly Val Asp Val Tyr Ala Val Ser Thr Asp Thr His Phe Thr His Lys
 65           70           75           80

Ala Trp His Ser Ser Ser Glu Thr Ile Ala Lys Ile Lys Tyr Ala Met
          85           90           95

Ile Gly Asp Pro Thr Gly Ala Leu Thr Arg Asn Phe Asp Asn Met Arg
          100          105          110

Glu Asp Glu Gly Leu Ala Asp Arg Ala Thr Phe Val Val Asp Pro Gln
          115          120          125

Gly Ile Ile Gln Ala Ile Glu Val Thr Ala Glu Gly Ile Gly Arg Asp
          130          135          140

Ala Ser Asp Leu Leu Arg Lys Ile Lys Ala Ala Gln Tyr Val Ala Ser
          145          150          155          160

His Pro Gly Glu Val Cys Pro Ala Lys Trp Lys Glu Gly Glu Ala Thr
          165          170          175

Leu Ala Pro Ser Leu Asp Leu Val Gly Lys Ile
          180          185

```

<210> 23

<211> 1486

<212> DNA

<213> Escherichia coli

<220>

<221> CDS

<222> (195)..(761)

<400> 23

```

aagggtagtt cagattacac ggtcacctgg aaagggggcc attttacttt ttatgcgctg 60

gcggtgcaaa gttcacaag ttgtcttacg aagggtgtaa ggtaaaactt atcgatttga 120

taatggaaac gcattaccgg aatcggcaaa aattgggttac cttacatctc atcgaaaaca 180

cggaggaagt atag atg tcc ttg att aac acc aaa att aaa cct ttt aaa 230
      Met Ser Leu Ile Asn Thr Lys Ile Lys Pro Phe Lys
        1           5           10

```


aac cag gca ttc aaa aac ggc gaa ttc atc gaa atc acc gaa aaa gat 278
 Asn Gln Ala Phe Lys Asn Gly Glu Phe Ile Glu Ile Thr Glu Lys Asp
 15 20 25

acc gaa ggc cgc tgg agc gtc ttc ttc ttc ttc tac ccg gct gac ttt 326
 Thr Glu Gly Arg Trp Ser Val Phe Phe Phe Phe Tyr Pro Ala Asp Phe
 30 35 40

act ttc gta tgc ccg acc gaa ctg ggt gac gtt gct gac cac tac gaa 374
 Thr Phe Val Cys Pro Thr Glu Leu Gly Asp Val Ala Asp His Tyr Glu
 45 50 55 60

gaa ctg cag aaa ctg ggc gta gac gta tac gca gta tct acc gat act 422
 Glu Leu Gln Lys Leu Gly Val Asp Val Tyr Ala Val Ser Thr Asp Thr
 65 70 75

cac ttc acc cac aaa gca tgg cac agc agc tct gaa acc atc gct aaa 470
 His Phe Thr His Lys Ala Trp His Ser Ser Ser Glu Thr Ile Ala Lys
 80 85 90

atc aaa tat gcg atg atc ggc gac ccg act ggc gcc ctg acc cgt aac 518
 Ile Lys Tyr Ala Met Ile Gly Asp Pro Thr Gly Ala Leu Thr Arg Asn
 95 100 105

ttc gac aac atg cgt gaa gat gaa ggt ctg gct gac cgt gcg acc ttc 566
 Phe Asp Asn Met Arg Glu Asp Glu Gly Leu Ala Asp Arg Ala Thr Phe
 110 115 120

gtt gtt gac ccg cag ggt atc atc cag gca atc gaa gtt acc gct gaa 614
 Val Val Asp Pro Gln Gly Ile Ile Gln Ala Ile Glu Val Thr Ala Glu
 125 130 135 140

ggc att ggc cgt gac gcg tct gac ctg ctg cgt aaa atc aaa gca gca 662
 Gly Ile Gly Arg Asp Ala Ser Asp Leu Leu Arg Lys Ile Lys Ala Ala
 145 150 155

cag tac gta gct tct cac cca ggt gaa gtt tgc ccg gct aaa tgg aaa 710
 Gln Tyr Val Ala Ser His Pro Gly Glu Val Cys Pro Ala Lys Trp Lys
 160 165 170

gaa ggt gaa gca act ctg gct ccg tct ctg gac ctg gtt ggt aaa atc 758
 Glu Gly Glu Ala Thr Leu Ala Pro Ser Leu Asp Leu Val Gly Lys Ile
 175 180 185

taa atttccttag tctttcacgc atagcggcgt tgcgtcgccc gctcaccg 811

tcacttactt gtgtaagctc ccggggattc acagctagcg ccttgctctg acgcgaaata 871

cttcggaat tcacctaat cttcgggtgc tgcggcgcat tttcttcccc gcaccatgat 931

gcaagctgca tccaggtagc cgcagaggcc gcttgcatga tgatgtttaa gagcccagga 991

gataaacatg ctcgacacaa atatgaaaac tcaactcaag gcttaccttg agaaattgac 1051

caagcctgtt gagttaattg ccacgctgga tgacagcgct aaatcggcag aaatcaagga 1111

actgttggct gaaatcgag aactgtcaga caaagtcacc tttaaagaag ataacagctt 1171

gccggtgcgt aagccgtctt tcctgatcac caaccaggt tccaaccagg ggccacgttt 1231
 tgcaggctct ccgctgggcc acgagttcac ctgctggta ctggcggtgc tgtggaccgg 1291
 tggcatccg tcgaaagaag cgcagtctct gctggagcag attcgccata ttgacgggtga 1351
 ttttgaattc gaaacctatt actcgctctc ttgccacaac tgcccggacg tgggtgcaggc 1411
 gctgaacctg atgagcgtac tgaaccgcg catcaagcac actgcaattg acggcggcac 1471
 cttccagaac gaaat 1486

<210> 24

<211> 188

<212> PRT

<213> Escherichia coli

<400> 24

Met	Ser	Leu	Ile	Asn	Thr	Lys	Ile	Lys	Pro	Phe	Lys	Asn	Gln	Ala	Phe	1	5	10	15
Lys	Asn	Gly	Glu	Phe	Ile	Glu	Ile	Thr	Glu	Lys	Asp	Thr	Glu	Gly	Arg	20	25	30	
Trp	Ser	Val	Phe	Phe	Phe	Phe	Tyr	Pro	Ala	Asp	Phe	Thr	Phe	Val	Cys	35	40	45	
Pro	Thr	Glu	Leu	Gly	Asp	Val	Ala	Asp	His	Tyr	Glu	Glu	Leu	Gln	Lys	50	55	60	
Leu	Gly	Val	Asp	Val	Tyr	Ala	Val	Ser	Thr	Asp	Thr	His	Phe	Thr	His	65	70	75	80
Lys	Ala	Trp	His	Ser	Ser	Ser	Glu	Thr	Ile	Ala	Lys	Ile	Lys	Tyr	Ala	85	90	95	
Met	Ile	Gly	Asp	Pro	Thr	Gly	Ala	Leu	Thr	Arg	Asn	Phe	Asp	Asn	Met	100	105	110	
Arg	Glu	Asp	Glu	Gly	Leu	Ala	Asp	Arg	Ala	Thr	Phe	Val	Val	Asp	Pro	115	120	125	
Gln	Gly	Ile	Ile	Gln	Ala	Ile	Glu	Val	Thr	Ala	Glu	Gly	Ile	Gly	Arg	130	135	140	
Asp	Ala	Ser	Asp	Leu	Leu	Arg	Lys	Ile	Lys	Ala	Ala	Gln	Tyr	Val	Ala	145	150	155	160
Ser	His	Pro	Gly	Glu	Val	Cys	Pro	Ala	Lys	Trp	Lys	Glu	Gly	Glu	Ala	165	170	175	
Thr	Leu	Ala	Pro	Ser	Leu	Asp	Leu	Val	Gly	Lys	Ile	180	185						